Global Perspectives on Accounting Education

Volume 6, *Issue* 1 2009 *Article* 1

READING AND UNDERSTANDING ACADEMIC RESEARCH IN ACCOUNTING: A GUIDE FOR STUDENTS

Teresa P. Gordon* Jason C. Porter†

*University of Idaho

†University of Idaho

READING AND UNDERSTANDING ACADEMIC RESEARCH IN ACCOUNTING: A GUIDE FOR STUDENTS

Teresa P. Gordon and Jason C. Porter

Abstract

The ability to read and understand academic research can be an important tool for practitioners in an increasingly complex accounting and business environment. This guide was developed to introduce students to the world of academic research. It is not intended for PhD students or others who wish to perform academic research. Instead, the guide should make published academic research more accessible and less intimidating so that future practitioners will be able to read empirical research and profitably apply the relevant findings. The guide begins by examining the importance of academic research for practitioners in accounting and next reviews the basics of the research process. With that background in place, we then give some guidelines and helpful hints for reading and evaluating academic papers. This guide has been used for several years to introduce master's degree students to academic literature in an accounting theory class. After reading this guide and seeing a demonstration presentation by the professor, students have been able to successfully read and discuss research findings.

KEYWORDS: Understanding empirical research; supplemental readings; importance of academic research; incorporating academic research in classroom

Global Perspectives on Accounting Education
Volume 6, 2009, 25-45

READING AND UNDERSTANDING ACADEMIC RESEARCH IN ACCOUNTING: A GUIDE FOR STUDENTS

Teresa P. Gordon
College of Business and Economics
University of Idaho
Moscow, Idaho
USA

Jason C. Porter
College of Business and Economics
University of Idaho
Moscow, Idaho
USA

ABSTRACT

The ability to read and understand academic research can be an important tool for practitioners in an increasingly complex accounting and business environment. This guide was developed to introduce students to the world of academic research. It is not intended for PhD students or others who wish to perform academic research. Instead, the guide should make published academic research more accessible and less intimidating so that future practitioners will be able to read empirical research and profitably apply the relevant findings. The guide begins by examining the importance of academic research for practitioners in accounting and next reviews the basics of the research process. With that background in place, we then give some guidelines and helpful hints for reading and evaluating academic papers. This guide has been used for several years to introduce master's degree students to academic literature in an accounting theory class. After reading this guide and seeing a demonstration presentation by the professor, students have been able to successfully read and discuss research findings.

Key words: Understanding empirical research, supplemental readings, importance of academic research, incorporating academic research in classroom

INTRODUCTION

There has long been a communication gap between the work of academic researchers and that of practicing accountants (for example, see Sterling 1973 and Zeff 2003). For researchers, the result of this gap is frustration that carefully prepared contributions to the field are ignored by those who could most benefit from new knowledge. For practitioners, the result seems to be a mix of gratitude for the efforts of academics to prepare the next generation of accountants and amusement at the time 'wasted' on what they feel to be irrelevant papers that end up gathering dust on the shelves of the library (Leisenring and Johnson 1994). This guide is an attempt to bridge the gap between academics and practitioners by introducing undergraduate and master's level students to the academic literature.

We begin with a discussion of the important role academic research can play in accounting practice. We then examine the accounting research process, specifically how it follows the scientific method. Finally, we provide a set of steps and helpful hints for students and other non-academics who want to quickly, rather than thoroughly, read and evaluate academic papers.

In the next section we look at the importance of academic research to you and your career and discuss how academic research is performed. With this general background on the research process, we then present a systematic approach designed to quickly get useful knowledge from academic papers. As a result, you will learn how to read and how to evaluate academic papers with confidence.

THE IMPORTANCE OF ACADEMIC RESEARCH

Practicing accountants, whether in public practice or industry, spend considerable time and effort conducting research for their clients. They have to decide how to implement new accounting or auditing standards, how to present unusual economic transactions in the financial statements, and how new tax laws impact their clients or employers. This research focuses on solving immediate problems for a single client or small group of clients. However, there is another type of accounting research that focuses on how the accounting profession affects the capital markets through academic accounting research. Academic accounting research looks at various topics in financial reporting, auditing, systems implementation, tax reporting, and other key issues from a scientific perspective. The studies use evidence from many different sources, including financial statements, stock prices, surveys, experiments, and even computer simulations and mathematical proofs. Research topics range from immediately useful aids to improving current audit procedures to big picture issues regarding the future direction of the profession. In addition, many papers focus on either the production or the use of accounting information. To put it another way, academic research looks at how accounting affects the world around us and how the world affects accounting. As a result, it can provide powerful information and insights for regulators, auditors, tax consultants, and other practicing accountants.

The academic research literature addresses all aspects of the accounting profession, from managerial accounting to analyst forecasts of earnings per share. One of accounting researchers' primary goals has been to examine the effectiveness of current accounting practices in conveying information to stakeholders (e. g., Guenther and Young, 2000). Researchers have addressed all aspects of this process, from the usefulness of managerial accounting methods (Lipe and Salterio,

2000) to the success of new audit methods (Bamber and Ramsay, 2000). These studies can give new insights to practitioners and regulators, especially when the evidence suggests that current methods are not as effective as they could be. In addition, they can improve the understanding of how stakeholders actually use the information accountants provide. Such studies can help practitioners find ways to produce information that is more useful. Other examples of major research topics covered by the academic literature include: how well current auditing techniques work and how to improve them; how tax laws affect companies' planning and accounting presentation; how managerial accounting methods help firms improve their use of the available information; how accounting information affects promotion and employment within firms; and how the change to IFRS will affect accounting and the capital markets. In addition to the practice of accounting, academics also conduct studies that test various methods for effectively teaching accounting topics.

Academic research plays a critical role in the creation of new knowledge. Although some would argue that this role is largely confined to the hard sciences (e.g., physics, chemistry, and biology), academic research also plays a part in developing our professional knowledge and practice of accounting. Academic papers address nearly every aspect of the accounting profession and provide insights that can aid in creating new auditing and accounting standards, improving practice, and identifying other important issues. Unfortunately, most practicing accountants have no training in reading academic research, which leads many to dismiss what could be very helpful information as either too complicated or too disconnected to be useful. After a little training, however, both current and future practitioners can start to see just how relevant and interesting academic research can be to accounting practice.

Wilks and Zimbelman's (2004) study is one example of research that has immediate relevance to practice. Earlier auditing standards provided numerous examples of fraud risk factors, but SAS No. 99 was the first to organize the factors into the fraud triangle. Regulators had expressed concerns that auditors were putting too much emphasis on their assessment of managers' attitudes and not enough emphasis on the opportunities and incentives to commit fraud. Wilks and Zimbelman examined whether first assessing each side of the fraud triangle and then combining those assessments would result in a better fraud risk assessment than making one holistic judgment. They found this twostep approach helpful in a low fraud risk setting, but not in a high risk setting. This finding suggests that using the fraud triangle may not improve fraud risk assessment in the most sensitive situations. It also demonstrates how academic research can help identify or test potentially useful auditing techniques.

Some academic research has a much broader focus, with implications for the future direction of the accounting profession. This type of research is just as valuable in the long-run but might not lead to immediate changes. Lev and Zarowin (1999), for example, conducted a study on stock market returns that suggested that financial statement information was becoming less important to investors; a finding that has considerable implications for the foundation of the profession. The authors suggested that the problem might be due to slow changes in accounting methods relative to changes in business practices and made suggestions about ways that the profession could improve.

Other studies look at narrower aspects of financial reporting such as managements' announcement of earnings. Bagnoli et al. (2003) tested whether the date that earnings are announced impacts a firm's stock price. They examined a group of firms where management voluntarily publicized the date of their upcoming earnings announcement. If an earnings announcement occurred

at a later date, investors appeared to become suspicious that managers had something to hide, and the stock price dropped. For every day the earnings announcement was late, earnings per share dropped by about a penny. The researchers also showed that investors' guesses are usually correct. While these results do not have a specific implication for regulators, they do provide an interesting insight into how investors react to the information accountants provide. It also seems to contradict the findings of Lev and Zarowin (1999), since these investors cared enough about accounting information to penalize a company when that information is not available when they expect it.

Another specific research area deals with the global movement toward the use of International Financial Reporting Standards (IFRS). This research examines both the standards themselves and the impact of implementation. For example, Leuz and Verrecchia (2000) looked at German firms where managers voluntarily reported financial statements following U.S. GAAP or IFRS. The evidence suggests that these companies were providing valuable additional information to the market, beyond that provided by statements prepared under German GAAP. However, the study revealed little evidence of differences between the new information provided by U.S. GAAP and the new information provided by IFRS. Although this study dealt with only a small sample of firms, the results support arguments being made for moving the U.S. toward adoption of, or convergence with, IFRS.

Many research studies have focused on earning manipulations, especially after recent scandals such as Enron, Tyco, and WorldCom. These studies have not only attempted to define what actually constitutes earnings manipulations, but have also addressed how to recognize manipulations, how to deter managers from manipulating their earnings numbers, and what other aspects of the financial statements might be affected. For example, Erickson et al. (2004) examined the tax effects of earnings manipulations and found that managers using manipulations to raise their earnings are willing to pay actual taxes on their fictitious earnings. This can result in a double loss to investors since they lose out on potential dividends or growth with the money that is paid in taxes as well as potential losses from company failure or legal fees and penalties when the manipulations are discovered.

The knowledge generated through academic research can provide valuable insights to aid regulators in the creation of new GAAP and auditing standards, auditors in their assurance work, and financial statement preparers in avoiding common pitfalls and the appearance of manipulating earnings. The findings also provide insight into the risks and issues facing the accounting profession. Although academic research provides important information, it often produces results more slowly than practitioners would like. In addition, the complex methods and writing style used by academics often hide many of the potential benefits. These weaknesses often turn away practitioners, the very people academics would most like to help. We hope to provide a few basics about academic methods and writing styles that will help you to read, evaluate and put to use the new information produced by academic accounting research. However, these are not the only benefits. Each of us is constantly bombarded by statistics that range from approval ratings for politicians to crime and divorce rates. The results may be trivial (which athlete has the best record) or serious (side effects of medication). One way or another, the tools and issues discussed in this paper can help you develop a working knowledge of research that will enable you to gather useful information from published papers and from every day statistics. More specifically, we hope to provide a basic introduction to research methodologies that will help you gather useful information from academic papers.

CREATING NEW KNOWLEDGE - THE RESEARCH PROCESS

The Scientific Method

Advances in science are most often the result of a process called the scientific method (Figure 1). The process begins when a researcher observes a set of events and develops a theory, or explanation, of what might be causing those events. The researcher then tests the theory to see how well it explains observed events. To do this, he or she uses the theory to develop a hypothesis, or prediction, about what will happen in a particular situation. Next, he or she designs a set of tests to determine if the hypothesis is correct. Test results consistent with the prediction confirm the hypothesis and, thus, the theory. When the results are inconsistent with the hypothesis, the theory is 'disconfirmed.' The researcher must then decide whether the theory needs some fine-tuning or if it should be replaced by a different theory.

Because it is difficult to create one explanation that covers all of the observed events, the creation of a theory is usually a long and time-consuming process. The first step is usually the observation of an anomaly that existing theories do not explain or the observation of a pattern in what was thought to be a random set of events. While the researcher might spend some time trying to fit these new patterns into existing theories, true discovery really involves thinking about the pattern and what might be causing that pattern. However, it is not always necessary to start from scratch. Accounting researchers often use existing theories from psychology, economics, and other fields, with some minor adjustments, as the basis for new accounting theories.

The theory of earnings manipulation provides a good example of this adaptive process. The theory that managers will manipulate earnings to mislead investors is based on a well-established finance theory called agency theory. Agency theory suggests that managers (agents) act in their own self-interest and put their own goals ahead of the owners' goals. Since the owners (principals) are not in a position to observe managers' actions, the resulting information asymmetry (agents knowing more than the owners) can be used to explain why techniques like audits and stock compensation plans are useful and why earnings manipulations occur. Although agency theory is a very broad, well-accepted theory, technically speaking, it is not an original theory either. It is derived from an economic theory called utility theory. Utility theory suggests that everyone wants to consume as much as they can for the lowest possible cost. In a way, agency theory is just one aspect of utility theory and the theory of earnings manipulations is one specific aspect of agency theory.

Background Research

The researcher's first step after becoming interested in a theory is to take a real or virtual trip to the library to find out what aspects of the theory have already been tested. After all, it does not make sense to spend time re-inventing the wheel. If others have already tested a hypothesis, it is usually better to move on and test a new aspect of the theory. Library research will usually provide other information as well, such as whether alternative theories that make opposing predictions exist, and what types of tests have been used to investigate this and similar theories in the past. Perhaps most importantly, the researcher can also find out if the prior research has been consistent. In other words, does all of the published evidence support the theory or is there some disagreement? All of this information helps the researcher develop both a stronger hypothesis and a stronger set of tests for that hypothesis.

If a researcher wants to study earnings management, a quick trip to the library would show that it has been the focus of many research studies. First, he or she would find that the original theory, agency theory, was developed in the finance literature by Jensen and Meckling (1976). After being used extensively in the finance literature, it was applied in accounting in various ways, including earnings management (Schipper, 1989). Next, the researcher would find many different papers examining how managers manipulate earnings, the appearance of the balance sheet, and other aspects of financial reporting intended to misrepresent performance (Kothari, 2001). To be interesting and useful, a new study would need to take into account all of the earlier research. For a topic like earnings management, the researcher may use tests similar to earlier studies but use a new time period, data source or perhaps a new statistical technique.

Developing a Hypothesis

After examining the existing theories, hypotheses and tests in the literature, the researcher can then use the theory he or she has chosen to create a specific hypothesis. A good hypothesis will focus on a relatively untested aspect of the theory and make a specific prediction about what will happen when a specific set of conditions exist. Hirst and Hopkins (1998) illustrate this process. U.S. GAAP currently allows companies to present their comprehensive income information at the end of the Income Statement, in a separate Statement of Comprehensive Income, or as an addition to the Statement of Stockholders' equity.[1] Although the three formats all present the same information, Hirst and Hopkins observed that most U.S. companies choose to put the information in the Statement of Stockholders' Equity. After observing that the comprehensive income information is usually negative and that investors tend to focus their attention on the bottom line income numbers (Sloan, 1996), Hirst and Hopkins hypothesized that firms put comprehensive income in the Statement of Stockholders' Equity to hide the information from investors, a form of financial statement manipulation. Here is a hypothesis from their paper (Hirst and Hopkins, 1998. p. 58):

> H1: The difference in analysts' stock price judgments when they value firms that do versus do not manage earnings will decrease as the clarity of disclosure of comprehensive income and its components increases.

At first glance, this seems complicated, but it is actually an example of a good hypothesis. First, it gives specific details of the sample they will actually test (financial analysts), what will be tested (analysts' stock price estimates for each format), and what the researchers expect to find (a smaller difference between estimated and actual stock prices when comprehensive income is presented more clearly). In addition, this hypothesis addresses an area of earnings management that had not been studied before (the comprehensive income disclosure format) and indicates its relationship to earnings management theory by suggesting that managers can hide negative information from their investors when the comprehensive income information is less transparent. Third, the hypothesis is general in scope since it applies to any type of firm. Fourth, the hypothesis

[1] The IASB began requiring a statement of comprehensive income in 2009 (see *International Accounting Standard 1, Presentation of Financial Statements*).

is stated clearly. A good hypothesis will aim for the 'elegance of simplicity.' In other words, it should avoid overloading the reader with lots of extra detail and specific facts.

Finally, the topic of the hypothesis should be of interest, both to the researcher and to those who will read the study. In this case, the hypothesis is interesting to academics because it gives additional information about the forms of earnings management, and to auditors and investors because it provides a potential indicator that firms are trying to hide information. It should also be interesting to chief financial officers (CFOs) and controllers because it provides a way to signal that a firm is not hiding anything through choice of format. Finally, it is interesting to regulators, such as the IASB, because they do not want to develop standards that allow companies to hide useful information from their stakeholders. In this case, Hirst and Hopkins' results that investors do not use comprehensive income information correctly when it is placed in the Statement of Stockholders' Equity might have influenced the IASB's standard on comprehensive income, since their new standard only allows the two income statement formats (see IAS 1, ¶81).

Designing the Tests

Once the researcher has determined the hypothesis, he or she can start identifying data sources and developing appropriate tests to examine that hypothesis. This process is usually one of the most time-consuming parts of doing research, since only a carefully constructed test will be empirically valid. Validity refers to how well the test actually addresses the research question, and ensuring validity is the most important part of designing the tests.

Internal and External Validity

Internal validity, at the purest level, refers to how well the study captures a cause-and-effect relationship. For example, does presentation format cause analysts to make forecast errors in Hirst and Hopkins' (1998) study? Because of the large number of alternative factors that exist in real situations, only an experiment can be used to test a cause-and-effect relationship. An experiment sets up a carefully controlled situation that ensures that the cause being tested, and only that cause, influences the effect. Hirst and Hopkins asked a group of financial analysts to make a stock price judgment on a fictitious company using only their own experience and the information provided by the researchers. Since the analysts were randomly assigned the format of comprehensive income they received and no other information was available to any of the analysts, differences in stock price judgments would only be caused by the comprehensive income format provided by the researchers. In experiments, academics refer to the different levels of the cause being tested (such as putting the comprehensive income information in a Statement of Comprehensive Income vs. a Statement of Stockholder's Equity) as manipulations or treatments. Other than the planned manipulations, everything experimental participants see and hear is exactly the same.

In a more general sense, internal validity refers to how well the study tests the relationship between events described by the hypothesis. All good studies, not just experiments, have some internal validity. Some studies look for relationships between variables instead of cause-and-effect relationships. In other words, they look for whether or not two variables covary or move together. These studies must still have some internal validity, but they cannot show a true cause-and-effect relationship. Still other studies test for differences between groups to understand (but not necessarily explain) what conditions exist in the real world, providing a foundation for later theory building.

Almost all research studies can be classified as either "true experiments," or as non-experiments or quasi-experiments.

Non-experiments are simple examinations of what currently exists. While they do not allow much control for internal validity, these discovery-type studies often provide some of the most interesting information. The simplest example of a non-experimental design is a survey. For example, a researcher could design a survey to ask a group of financial executives, auditors, and financial analysts whether the presentation of comprehensive income signals earnings management. This survey would provide an interesting look at what professionals think about the presentation and usefulness of comprehensive income. However, with a survey there is no way of ensuring that the answers are true.[2] A recent study by Nelson et al. (2003) provides a good example of an interesting non-experiment. The researchers surveyed auditors about the types of earnings manipulations they had observed in actual financial statements. While the study does not address the causes of the manipulations or how the auditor found those manipulations, the list of documented manipulations can have important implications for regulators setting GAAP and auditing standards, for auditors considering what aspects of the financial statements to test for possible manipulation, and for firms wanting to avoid the appearance of manipulating earnings.

Quasi-experiments, on the other hand, provide some of the control of an experiment while still retaining the real world power of a non-experiment. Academics also refer to quasi-experiments as "natural experiments," since they occur when a group of individuals or companies self-select into different groups. Because the researcher does not control the selection process, the reader cannot be sure that some other event did not cause the choice that made the groups differ. This limits the internal validity of the study. In the comprehensive income example, managers decide for themselves how to present comprehensive income. By making the choice, companies are naturally grouped into the different manipulation levels that Hirst and Hopkins (1998) artificially introduced in their experiment. In fact, a quasi-experiment by Lee et al. (2006) examines the comprehensive income format choice of a sample of publicly traded insurance firms. The researchers find evidence that insurers with a tendency to manage earnings in other ways or that have poor disclosure quality are more likely to put their comprehensive income information in the statement of stockholders' equity. With a quasi-experiment, the researchers could not be sure that the choice of format was intended as earnings management. However, the finding still suggests that regulators might want to consider requiring a standard format.

Because non-experiments and quasi-experiments use real world data, they tend to have higher levels of external validity. External validity refers to how well the results from a study can be applied to other settings, such as a specific client or to other investors. While most researchers agree that internal validity is the most important aspect of a study, external validity runs a close second. If the cause-and-effect relationship occurs only in a laboratory, it may be interesting, but not really

[2] With a simple multiple-choice question like 'Which of the following presentations does your company use for comprehensive income?' there is not much risk that survey participants would lie. However, if asked why their companies use a particular method, they might not be in a position to know the real reason or they might choose to provide a socially or professionally acceptable answer rather than the real one. In other situations, such as asking whether or not a firm manipulates its earnings, it is possible or even likely that the managers will lie or not respond to the survey.

important to practice. For an applied science, like accounting, external validity makes the results of a study useful. Thus, a good study needs to have internal validity to show that the relationship being tested really occurs and external validity to show that the relationship being tested occurs in natural or real world settings.

After defining his or her hypothesis, the researcher needs to decide which type of study to perform. In accounting, a quasi-experiment is probably the most frequent research design, although 'true' experiments and non-experiments are also used. Using a quasi-experiment, however, allows the researcher to achieve acceptable levels of internal validity to satisfy the academic audience and high enough levels of external validity to ensure that the findings are applicable to practice. The choice for a particular study is usually a tradeoff between internal and external validity goals for the hypothesis. Over time, researchers may use all three of the research designs to study a theory. Because of the tradeoff between internal and external validity in each category, it is only through this combination of evidence from all three designs that a theory can be confirmed or disconfirmed. Figure 2 provides a flowchart identifying the different categories of research design.

Construct Validity and Defining Variables

Once the researcher has determined the category of study that best matches the hypothesis, he or she must define the events and conditions that will be measured. This leads to another validity issue: construct validity. Construct validity refers to how well the variables used in a study capture the ideas and events in the hypothesis. Variables are the events and conditions that will actually be measured in a study. In some cases, the variables can be easily observed, such as net income or dividends. In other situations, the variables can be almost impossible to observe, such as managers' intentions or an individual's natural ability. In a given study, the event of interest is known as the dependent variable and the cause being tested as the independent variable. Most studies will also include a number of control variables – other variables that may be associated with alternative explanations for changes in the dependent variable. The careful use of control variables can greatly improve the study's internal validity. Examples of control variables include the size of a firm, the composition of the board of directors, the type of audit firm (large or small) that performs the audit, and risk characteristics of the industry.

In the example of an experiment described previously, Hirst and Hopkins wanted to test how the comprehensive income disclosure format (the independent variable) affected analysts' stock price estimates (the dependent variable). However, the comprehensive income disclosure format is not the only thing that affects a stock price estimate. Some of the possible control variables include the rest of the financial statement information, the strength of the economy, the performance of competitors, new rules and regulations, news reports, auditor reports, and the management's letters. In a quasi-experiment, the researcher adds as many control variables as he or she can, but it is impossible to control for everything. By running an experiment, however, Hirst and Hopkins were able to control all of the information that their participants received, thereby eliminating the alternative effects previously mentioned, as well as many others, without having to include large numbers of control variables in their statistical tests. In contrast, the quasi-experiment by Lee et al. (2006) had to include a number of control variables such as the size of the firm, profitability, use of an international auditing firm (e.g. KPMG or PWC), the volatility of comprehensive income, a financial quality rating, the number of analysts following the firm, and the daily bid-ask spread. And even with all of

these control variables, they still could not be sure that the choice of comprehensive income format was a form of earnings management. It could have just been easier for the accountant or they could have flipped a coin. No matter how involved the statistical method, the researcher using a quasi-experiment can never be sure that an important variable was not omitted.

While all researchers must carefully consider what variables will be included in their study, researchers using hard to obtain or hard to measure variables must be especially careful. Researchers often use surrogate variables to substitute for hard-to-collect data. For example, debt covenants may prevent certain actions by managers such as paying dividends. It is both time consuming and expensive to examine all of a company's debt instruments, to identify a list of debt covenants, and then to check whether any of those covenants have been violated. Instead, a researcher might choose to use the debt to equity ratio as a surrogate to indicate how close a company is to violating its debt covenants, assuming that companies with high debt to equity ratios are closer to violating their covenants than companies with smaller ratios. When a surrogate variable is used to measure something else, the logic behind the selection should be sound and well documented by the researcher. In this case, the logic behind using the debt to equity ratio as a surrogate is that companies with little debt will probably not be close to violating any debt covenants they might have.

The most challenging situation for researchers is when the desired variable is not just hard-to-collect but unobservable (such as management intent or need for achievement). For example, a researcher might use a bonus contract as a surrogate measure for managers' intentions. Since managers probably want to get the largest bonus possible, the researcher assumes that their intention would be to maximize earnings resulting in the desired bonus. An alternative hypothesis might be that managers with a high need for achievement would also work hard to maximize earnings. The difficulty lies in the fact that managers have incentives (like taxes, audit results, or regulatory scrutiny) other than maximizing their bonus. Unobservable variables like intentions and personality are referred to as constructs, and the question is whether the chosen surrogates capture or measure the constructs appropriately. Evaluating construct validity is very challenging when it comes to unobservable variables. Researchers are sometimes reduced to making sure the surrogate seems plausible or has face validity. A study that otherwise has good internal and external validity will be less reliable if one or more variables lack construct validity.

Sample Selection

Once the researcher has decided on the type of study to perform and the variables to be included, he or she can begin gathering the necessary data. The first step in gathering data is to determine the population. The population is made up of all of the entities, firms, or individuals of interest to the current hypothesis. In the case of Hirst and Hopkins (1998), the population was comprised of all analysts working in U.S. stock markets. Usually, however, it is not practical to study the entire population.[3] For example, imagine how difficult it would be to get a survey completed by all of the investors in the U.S. stock markets or by all of the managers of all of the companies that follow U.S. GAAP. Since studying the full population is impractical in most cases, the researcher

[3] Some studies do allow researchers to gather the full population. For example, if an accounting firm wanted to learn the reaction of the partners to a proposed company policy, it would be possible to survey the full population.

will instead draw a sample, or subsection of the population. He or she will then perform the tests on the sample and use statistical assumptions to apply those results to the entire population.

While not as difficult as gathering the entire population, gathering a sample still requires careful thought and, usually, a great deal of effort. The best way to gather this sample is to use random selection. In this process, the researcher randomly selects individuals or firms from a list of the population. He or she will then try to convince those individuals or firms chosen to participate in the experiment or to provide the information needed. The objective of a random sample is to make sure that every observation in the population has an equal chance of being selected. When only a few individuals or firms from a sample agree to participate, it is possible that those who refused are systematically different from the participants. This possibility is called nonresponse bias. This problem is almost always mentioned as a limitation in survey research.

Although random samples are considered to be the best, nonrandom samples are often used. In many cases, researchers are forced to use nonrandom samples because the preparers, auditors, and users of accounting information have concerns about confidentiality or little incentive to spend time on a survey or experiment that does not offer any immediate rewards. Even the participants in experiments rarely come from a random sample of the population. For example, Hirst and Hopkins used a nonrandom or 'convenience' sample of financial analysts willing to participate in their study. While this sample was easier to get than a true random sample of all analysts, they came from only one group of analysts. Do results from this study, then, truly represent the population of all analysts? In other words, can the results be generalized to the population? Without a random sample, there is no way to know. The difficulty in getting people to participate in research studies is another reason why quasi-experiments, which pull data from existing historical databases, are popular in accounting. Unfortunately, using a nonrandom sample weakens the external validity of a study, since the sample will probably not represent the entire population.

The final decision in creating a sample is determining how many observations need to be gathered. To keep the costs of research down, the researcher will not want to spend time and money collecting more observations than necessary. However, if too few observations are collected, the researcher will be unable to run the statistical analyses. As a general rule, bigger samples are usually better, since they will give a better indication of how the population would react in the situation. On the other hand, some statistical tests get skewed if too many observations are used and almost anything is "statistically significant" even if the effect is very small. For most types of analysis, the necessary sample size can be determined using well-developed statistical techniques. A basic rule is at least ten observations are needed for each different independent or control variable with an absolute minimum of 100 observations (Cohen, 1990).

Finishing a Study

Once the researcher has developed the hypotheses, designed the tests, selected the sample, and collected the data, he or she is finally ready to perform the tests and examine the results. Statistical tests are used to compare the actual results from the collected data with the predicted results of the hypotheses. Typically, a study has one or two primary tests and several additional tests or statistical methods that are used to confirm the primary results. This process is called sensitivity analysis. The entire research process, from theory to results, is then written into a paper and submitted to an academic journal.

The journal editor will generally send the paper to two or more other researchers for a thorough review. The reviewers, usually experts in the area, will critically evaluate the paper. They will question everything the researcher has done, compare it to existing research, and challenge all of the design choices that were made. After reading and critiquing the paper and the research, the reviewers will then write a set of suggestions for improving the paper that are sent back to the editor. Using these reviews, the editor then decides to accept the paper with minor revisions, reject the paper, or to ask for a major revision and another submission. The author then gets a letter from the editor with the editor's decision and copies of the reviewers' comments.

Papers are rarely accepted as submitted. Instead, most researchers are required to respond to all of the reviewers' suggestions before the paper is published. When it is returned to the journal, the paper is usually subjected to the same lengthy review process a second time. Often a researcher must go through several iterations of this process before his or her study is finally published. Because of this lengthy process, the paper that is finally published is generally greatly improved from the original submission. Incidentally, neither the researchers nor the reviewers are paid for the work that they put into a project. The rewards are intangible, although they can have a tangible impact on the researcher's career, since many colleges and universities will only retain professors who successfully publish his or her research.

Summary

In this section we discussed the process used by academic researchers to create their papers. First, we discussed the scientific method and the creation of theories. Next, we discussed how a good researcher will develop his or her hypothesis after carefully reading the prior literature in the area. Then, we discussed how a researcher must decide on the category of study that best fits his or her hypothesis—an experiment, a non-experiment, or a quasi-experiment—and the tradeoffs between external and internal validity that must be carefully considered in making this decision. We also discussed the creation of the variables used in the study, the sample selection process, and the impact of those decisions on the internal, external, and construct validity of the paper. Finally, we discussed the process of finishing a paper and getting it published.

READING A RESEARCH PAPER

Now that we have discussed what is involved in the research process, we will provide techniques that can be used to efficiently and effectively obtain useful information from a published paper. Fortunately, research papers in accounting typically follow a standard pattern which makes it easy to find the most important aspects of a study and read them first. Each paper begins with an abstract and an introduction which summarize the important points of the paper. In addition, each paper ends with a conclusion or final comments section that reemphasizes those aspects of the study that the researcher feels are most important. Although readers tend to start at the beginning of an article and read straight through to the end, the best way to read an academic research paper is to start with the abstract, the introduction, and the conclusion sections. Table 1 provides a summary of our suggestions for reading an academic paper.

TABLE 1
Hints for Reading an Academic Paper

1. Read the abstract, introduction and conclusion to determine the question being asked and why that question is important.
2. Decide whether the question is interesting or important.
3. Make note of the important aspects of the paper (research question, method, etc.) for reference while reading the rest of the paper.
4. Read the literature review and examine the bibliography to determine the position of the paper.
5. Read the sample selection, method, and results sections of the paper to determine its validity:
   a. What category of study is it?
   b. If it is an experiment, what steps did the researcher take to improve external validity?
   c. If it is a quasi- or non-experiment, what steps did the researcher take to improve internal validity?
   d. Did the researcher carefully design the study to capitalize on the natural strengths of the category of study?
   e. Do the variables really capture the constructs the researcher claims they capture? Is there a more appropriate variable?
   f. What limitations does the researcher point out?
6. After assessing validity, consider the conclusions made by the researcher.
   a. Examine the results of the primary test. Are the results consistent with the researcher's explanation?
   b. Are the stated conclusions consistent with the results? Does the narrative discussion make unjustified claims about the usefulness of the study?

Getting an Overview

It may help to ask some basic questions and make notes of the answers to help understand the paper. First, what is the main question the paper is trying to answer and why does the researcher feel the answer to this question is important? One should also consider whether this question is interesting. Keep an open mind when making this decision. Most of us have a natural tendency to focus on our own area of accounting expertise, but we should not automatically discard papers in different areas. Papers will be interesting because of their potential impact on the profession, because they provide insights on an aspect of accounting that can be applied to other areas, or because they are simply interesting issues.

Assuming the paper seems to be useful or interesting, what do the authors have to say about their methods and the limitations to those methods in their introduction and conclusion? Since the remaining sections of the paper will be full of technical statistical terms, academic jargon, and discussions of the researcher's choices, it is useful to gather some basic details in these summary sections that can later be used as reference points. Make a note about anything that is of concern. For example, if you are not comfortable with an assumption the researcher has made or if you do not understand a point, write it down. While reading the remainder of the paper, try to evaluate the claims of the researcher. If the sample selection or variables being used do not make sense, then there

probably is something to wonder about. Just because a paper gets published does not mean that it is perfect. One researcher describes this process as 'a good sense of smell.' If something smells a bit rotten, it probably is. If the issue is not addressed satisfactorily in the paper, one can decide later whether to believe some of the conclusions of the paper or to find other papers on the topic

Locating the Paper on the Tree of Knowledge

With a good feel for the topic of the paper and a general outline of the research method, the next step is to get a sense of its position in the literature. This step will provide some important background that will help in judging the paper's conclusions, so it is usually worth at least a little time. There are two ways to get a feel for the position of a paper. The first is to look at the bibliography length. Longer bibliographies usually indicate a more mature theory with more developed tests and hypotheses. Shorter bibliographies, on the other hand, often indicate development of a newer theory. One can also get a feel for the position of the paper by reading through the literature review or hypothesis development section, which typically appears immediately after the introduction. Surprisingly, review sections with more detailed descriptions of the theory and the evidence surrounding it may indicate a less developed theory. A literature review consisting largely of lists of other studies is often a sign of a mature theory, as will be discussed below. As a general rule, newer theories usually have the broadest, most interesting questions for a general audience. However, they also have more validity issues. As a theory becomes more developed, the hypotheses become more refined. This leads to narrow questions that are more interesting for practitioners and researchers that focus in that area. While the appeal of the study might not be as broad, these papers usually have stronger internal and external validity. In addition, they usually use methods and constructs that have been refined and accepted.

One difficulty with papers investigating more established theories is the limited publication space in most academic journals. Because of this limitation, the literature review sections in these papers are often abridged to save space. They start to resemble lists of paper summaries rather than a true discussion of the prior literature. Since academic papers are typically written by researchers for other researchers, authors of papers based on older theories often assume that readers have read much of the prior research in the area and therefore have a working knowledge of the most common tests and constructs. If you find the topic interesting, you may need to read some of the earlier studies to get a feel for the questions that have already been asked, what the answers to those questions were, what tests were used, and what the terminology means.

Assessing Research Design and Validity

All of the preliminary reading that has been done up to this point provides an important foundation for the most important step in reading an academic paper: deciding whether or not to believe the researcher's conclusions. In other words, the final step is determining the validity of the paper. While this is the most important step, it will also be the most difficult. To effectively do this, it is usually best to think about the three aspects of validity discussed earlier: internal, external, and construct. Keep in mind that no research study will be perfect. It is the slow accumulation of research results that advances accounting knowledge and overcomes the weaknesses in any individual study.

For a beginning reader of academic research, it is probably easier to think about each of the validity issues individually. To do so, one should read the heart of the paper: the sample selection,

method, and analysis or results sections, which normally appear in that order after the literature review section. Before getting started, it is important to ensure that the question the researcher is trying to answer, the reason for asking the question, and what the author states as the conclusion are understood. Keep your notes handy for reference while reading. This will help to avoid getting lost in the technical details. Next, identify the type of study (experiment, non-experiment or quasi-experiment), if it has not already been determined.

When assessing the internal and external validity of the study, remember that it is difficult, if not impossible, to have high levels of both internal and external validity in one study. Experiments usually have high internal validity because of the level of control allowed but low external validity because of those same extensive controls. In Hirst and Hopkins (1998), for example, the participants in the experiment received a simplified set of financial statements with no notes. They did not have any information about prior years, other analysts' reports, past stock price movements, the economy, or the firms' competitors. This simplification was necessary for the researchers to get a clear picture of how comprehensive income format affects investors' decisions, but it is possible that their results will not hold in a real world setting. It could be that comprehensive income format is very important in the simplified setting of Hirst and Hopkins' experiment but does not matter to investors when they have the rest of their normal information set with which to work. If the study is an experiment, look for any aspects the researcher included to improve the external validity. These could include an exceptional group of participants who strongly represent the population, a task that closely matches a task the participants would normally perform, or even a small real world sample gathered in addition to the main experimental test. Researchers may also include a discussion of manipulation checks, but these are usually more important for evaluating internal validity.

Just because a study is an experiment does not mean that it has good internal validity. Consider all the possible causes for the event being studied. Then, while reading the researcher's description of the experiment, think about whether or not the experiment's design eliminates the alternative explanations. Keep in mind that many alternative causes, such as natural ability or having a good day, are eliminated by random assignment. Note that there is a difference between the random selection of a sample and random assignment. Random selection refers to the way the participants were selected to participate in the experiment and serves to improve external validity. Random assignment, on the other hand, refers to randomly assigning each participant to the manipulations of the study. According to statistical theory, the use of random assignment will eliminate unintended differences between groups, since each group should look much the same as the other groups. This leads to higher levels of internal validity and is one of the principal strengths of an experiment. In their study, Hirst and Hopkins (1998) randomly assigned their analyst participants to the three types of comprehensive income format. By doing so, they eliminated the effects of age, experience, raw talent, education, and other personal issues of the analysts since equal proportions of those qualities should be found in each of the manipulation groups.

While reading the description of the experiment, see if the test used a control group. A control group is a group of participants who do not receive any manipulations. In Hirst and Hopkins (1998), a control group would have been a group of analysts who assessed stock price using all of the information except comprehensive income, since that information would have been omitted from their packet. The use of a control group allows researchers to test whether the manipulations have an affect. If Hirst and Hopkins had found evidence that the stock price from a control group was the

same as the stock price for the other three groups (the three manipulations of the comprehensive income format), then the conclusion would have been that comprehensive income did not have any effect on analysts.

Quasi-experiments and non-experiments usually have high external validity because of their larger sample sizes and the use of actual events but low internal validity because it is difficult to show a true cause-and-effect relationship. If the study is a non-experiment or quasi-experiment, look for any aspects the researcher included to improve internal validity. These could include an extensive group of control variables, a reduced sample size to eliminate observations associated with the event for reasons other than the hypothesized cause, or from the use of several alternative statistical tests. Usually the sensitivity analysis, or additional testing, is typically found in a special section towards the end of the paper.

Just because quasi- and non-experiments use real-world data does not necessarily mean that they have high levels of external validity. As with an experiment, they must be performed correctly to capitalize on their natural strengths. Spend some time looking at the sample selection section. Since the most powerful aspect of these studies is their real world data, this section becomes the easiest way to judge how externally valid the study's results will be. Consider the following questions while reading through this section: Is the sample period appropriate for the question being tested? Does the sample seem like a good subset of the population in which the researcher claims to be interested? Are there enough observations in the sample to make the results applicable to the entire population? While considering these questions, keep in mind that studies drawing random samples from the population of interest have more external validity than studies using nonrandom samples.[4] As an example, Lee et al. (2006) examined the comprehensive income format of 82 publicly traded insurance firms from 1994-2001. While this study is interesting because it provides evidence on how actual firms decide which comprehensive income format to use, one should consider whether these firms are representative of the rest of the population of all firms following U.S. GAAP, especially when they all come from one industry. Also, the applicability of findings from before the end of the internet bubble to the financial markets today may be questioned. Finally, the sample size (82) may be suspect as not being sufficient. Because of these issues, it might be difficult to apply and use the results of this study, no matter how impressive the internal validity.

After considering the paper's internal and external validity, spend a little time considering the construct validity of the variables being used. Start by looking at the method section to see what measures are used to capture the dependent and independent variables, and to a lesser extent, the control variables. Consider whether each variable really captures what the researcher says it captures. In addition, consider whether there is of a better measure for the construct. If you do not agree with the variable they are using but cannot think of anything better, then the variable being used might be the best available at the moment. In this case, accept the results with a grain of salt and keep looking for a future study that develops and uses a better measure.

---

[4]One example of nonrandom sampling that can be more interesting than random sampling is to deliberately seek out extreme or unusual cases. For example, firms that have been or are under investigation by the SEC for earnings management might be a more interesting group to test for hiding information in comprehensive income. While this directed sampling approach usually provides stronger evidence of the hypothesis being tested, it also limits the external validity of the study.

Most research papers will have a section or paragraph at the end of the paper that discusses limitations.[5] This is a very useful guide to potential problem areas. Ideally, you will have already thought of many of the problems the authors highlight. While reading this section, consider whether the researcher discusses potential weaknesses openly. Does he or she provide reasons why the validity issues might not affect the results of the paper? Does it appear that the researcher has given the weaknesses careful thought? This last is the most important question. If the researcher has spent significant time and energy thinking about and discussing the limitations of the study, it is usually an indication that he or she is a careful researcher. It is likely that what is presented will be quality work, despite the weaknesses. You can give more credence to the conclusions and results from this type of paper than a paper that tries to cover up the weaknesses or dismiss them out of hand.

Evaluating the Conclusions

The final step in reading an academic research paper is to consider the researcher's conclusions based on an assessment of the paper's validity. To do that, start by examining the results of the principal tests. For many students, this is the most difficult part of the assessment. Other than audit sampling, statistical tests are not commonly used in accounting classes or in the accounting profession, especially not the more complicated tests commonly used in academia. Luckily, most papers' primary test is a basic, simple version of the more complicated tests that form the majority of the paper. In this case, you can assess that primary test, ensure that the sensitivity analysis results are consistent with the primary results (usually through a quick glance at the tables), and be satisfied. If you feel uncomfortable assessing the statistical tests (or other parts of the paper), rely on the fact that the statistical analysis was evaluated and approved by at least two other academics. While you should still evaluate the paper, keep in mind that these reviewers were probably chosen because of their experience in both the research area and in the statistical techniques used.

Even if you rely on the expertise of the reviewers, skim or read the statistics presented in the paper. Statistics provide a wealth of information about a study. Summary statistics should statistics are provided by most studies to describe their sample. These statistics are commonly found in the first or second table of the paper and should contain basic statistical information such as the mean or average value and the standard deviation of the variables of interest. The authors may also provide the median or middle ranked value with the range from minimum to maximum. The standard deviation is a measure of the range of values reported for each variable. Small values indicate that the variable is relatively the same for all of the observations in the study, while large values indicate that the variable is widely different from one observation to the next. In most cases, the summary statistics will provide a feel for the sample and for the variables being used in the tests. Watch for variables with large differences between the mean and median values, and for variables with large standard deviations. While this does not necessarily mean there is a problem with the sample, it usually means that some special statistical tests need to be performed to ensure that the primary test is appropriate.

---

[5] Remember to check the footnotes too. Helpful information that editors and reviewers wanted the authors to add regarding alternative independent variables, statistical tests, or potential weaknesses may be found. These details can help in evaluating the methods used in the paper.

The next step is to look at the primary test of the study. In most accounting papers, the primary test will be one of three things: a t-test, a simple regression, or a multivariate regression. A t-test is a simple comparison between the mean values of two samples. It is commonly used to test the differences between the groups in experiments, such as comparing the stock price estimate of the analysts from the three different comprehensive income formats in Hirst and Hopkins (1998). T-tests are also used to test the differences between naturally occurring groups in quasi-experiments and the differences between survey results from different groups of respondents in non-experiments. The results of a t-test are typically reported as either a number called a t-value or as a percentage called a p-value. In most cases, academics and statisticians consider two groups to be different only if the t-value is greater than 1.96 or the p-value is less than or equal to 0.05. These numbers can be interpreted as saying "there is a less than a 5 percent chance that the mean values of the two groups are really equal."

Regression tests are set up a little differently. In a regression, the researcher is looking for the covariance of the two variables.[6] Covariance refers to the way the two or more variables move together. Using a regression test, researchers can examine whether an increase in the independent variable is associated with an increase or a decrease in the dependent variable. In some cases, the hypothesis will predict that the independent variable is associated with an increase in the dependent variable (such as an increase in net income being associated with an increase in stock price). In others, the hypothesis will predict the opposite relationship (for example, an increase in size of the audit firm is associated with a decrease in earnings management). Simple regressions examine the relationship between the dependent variable and one independent variable. Multivariate regressions also include a series of control variables to ensure that the relation between the dependent and independent variables is not being caused by something else. Thus, a multivariate regression is usually associated with higher levels of internal validity.

The first thing to look at in a regression test is the adjusted R^2 value. This statistic summarizes the overall success of the analysis. The adjusted R^2 is interpreted as the percentage of the overall movement in the dependent variable that the independent and control variables explain. A model that explains just one or two percent of the movement of the dependent variable is not particularly impressive, even if the coefficients appear to be significant. Researchers also report whether the overall model is statistically significant by reporting the F-value or p-value. If the results of these tests are not significant ($p > 0.05$), then the entire model should probably be rejected.

If the adjusted R^2 is acceptable, the next thing to consider is the signed coefficient for each independent and control variable. The coefficient measures how a change in the independent variable affects the dependent variable. For example, assume that a researcher found a coefficient value of 1.5 when testing the relationship between cash flow from operations and net income. That would mean that for every $1 increase in cash flow from operations, net income would go up by $1.50. There should also be a t-statistic or p-value reported for each coefficient that can be interpreted in the same way as described for the results of a t-test (using the same cutoff values). If the reported results (usually in a table) are not significant, then the researcher hasn't found any relationship

[6] You might also see correlations mentioned in research papers. A correlation is another measure of association between two variables. Larger correlation coefficient values indicate a stronger relationship and the sign of the coefficient indicates whether the two variables move in the same (positive) or in the opposite direction (negative).

between that variable and the dependent variable. In other words, a change in one does not relate to a change in the other.[7]

In some cases, however, a researcher will report a seemingly significant t-value and p-value for a very small coefficient. Perhaps the best examples of this relationship are ERCs, or Earnings Response Coefficients. These coefficients come from tests of the relationship between earnings and stock price, and researchers typically report coefficient values of less than 0.05 while still reporting a significant t-value and p-value (e. g., Kormendi and Lipe, 1987; Easton and Zmijewski, 1989). As with a small adjusted R^2, the relationship is statistically significant, but there is little practical significance. Always evaluate regression results for both statistical and practical significance before accepting the researcher's conclusions.

While the statistical tests discussed above are the most common used in the accounting academic literature, they are not the only statistical tests available. However, it is not necessary to be an expert in statistics to evaluate a paper. Many statistical glossaries or online dictionaries can be used to look up new statistical terms. Several online dictionaries (such as www.dictionary.com) provide good descriptions of statistical terms and tests. Similarly, typing statistical terms into an internet search engine will typically bring up a professor's introductory lecture on the subject. A quick glance through the slides will often provide enough of an understanding to feel comfortable with the analysis used in the paper.

After examining the test results, the next step is to assess the paper's conclusions. If the paper has good internal validity, good external validity, sufficient construct validity and significant results (both statistically and practically), then you can have confidence that the conclusions of the paper are substantially correct. However, it is important to ensure that the researcher has not made stronger conclusions than his or her tests warrant. If the paper uses a small sample of only 50 firms from a single industry, then perhaps any implication that the results apply to all firms is not appropriate. Since everyone has a tendency to exaggerate the importance of their own work, authors may go a little too far in their claims. When in doubt, go with "the facts" reported in the results section rather than the narrative conclusions that may claim a bit too much. If the author seems to be making conclusions that go way beyond his or her actual findings, it may be appropriate to reconsider the other validity issues as well.

Summary

In this section, some relatively painless ways to efficiently read through an academic paper were discussed. The most important advice was to start with the abstract, the introduction, and the conclusion sections. These sections provide a good feel for what questions the paper addresses, the basic way the researcher tested the question, and what he or she concludes based on the results. We then described ways to assess the validity of the paper and to evaluate the conclusions the researcher has made.

Much of the discussion focused on the weaknesses and problems with academic research. While this focus was necessary for evaluating academic research papers, it tends to make a reader too skeptical. Remember, no study will be perfect. Each paper tries to move existing knowledge

[7] Even when the coefficient is relatively large and statistically significant, the model may have little or no practical impact if the adjusted R^2 is very small. This often happens with very large sample sizes.

forward by just one step, not to answer all of the remaining questions about an issue. Well written papers use as many tests as possible, make careful claims based on the reported findings, and openly acknowledge and discuss potential weaknesses. The results of these papers can be accepted and used even though no paper is perfect. In this way, academic research is similar to audited financial statements: materially correct, not perfect.

CONCLUSION

In an ideal world, academic research articles would be accompanied by adaptations for students and practitioners. Such adaptations are a useful contribution and are slowly becoming more common. The FASB has even distributed some "long abstracts" of research papers they feel to be relevant to standard setting. Unfortunately, few research papers come with such a summary. Most research is done by professors in colleges and universities whose careers depend on publishing papers in the most respected journals in their field. Such journals have a limited audience (almost entirely researchers), and the papers they publish are written for that audience.

Given the differences between academic and professional writing, it has traditionally been difficult for practitioners to get much value out of current academic research. Most practitioners feel that academics are wasting their time with topics that do not matter in the real world and most academics feel that practitioners do not understand their work. The goal of this paper has been to bring the two groups closer together by providing future practitioners, namely current accounting students, with enough background that they can read academic research. This paper has provided an overview of the research process and tips for reading published papers with confidence. We strongly believe that the combination of practical experience, sound theories, and empirical results will, in the long run, benefit academics, practitioners, and students alike.

TEACHING NOTES

Teaching notes are available from the editor. Send a request from the "For Contributors" page of the journal website, http://gpae.bryant.edu.

REFERENCES AND SOURCES

Bagnoli, M., W. Kross, and S. Watts. 2002. The Information in Management's Expected Earnings Report Date: A Day Late, a Penny Short. *Journal of Accounting Research* (Vol. 40) 1275-1296.

Bamber, E. M., and R. J. Ramsay. 2000. The Effects of Specialization in Audit Workpaper Review on Review Efficiency and Reviewers' Confidence. *Auditing* (Vol. 19) 147-157.

Cohen, J. 1990. Things I Have Learned (So Far). *American Psychologist* (Vol. 45) 1304-1312.

Easton, P. D., and M. E. Zmijewski. 1989. Cross-Sectional Variation in the Stock Market Response to Accounting Earnings Announcements. *Journal of Accounting and Economics* (Vol. 11) 117-141.

Erickson, M., M. Hanlon, and E. L. Maydew. 2004. How Much Will Firms Pay for Earnings That Do Not Exist? Evidence of Taxes Paid on Allegedly Fraudulent Earnings. *Accounting Review* (Vol. 79) 387-408.

Guenther, D. A., and D. Young. 2000. The Association Between Financial Accounting Measures and Real Economic Activity: A Multinational Study. *Journal of Accounting and Economics* (Vol. 29) 55-72.

Hirst, D. E., and P. E. Hopkins. 1998. Comprehensive Income Reporting and Analysts' Valuation Judgments. *Journal of Accounting Research* (Vol. 36) 47-83.

International Accounting Standards Board. 2009. *Presentation of Financial Statements*. International Accounting Standard No. 1. (London, United Kingdom, IASB).

Jensen, M. C., and W. H. Meckling. 1976. Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics* (Vol. 3) 305-360.

Kormendi, R., and R. Lipe. 1987. Earnings Innovations, Earnings Persistence, and Stock Returns. *Journal of Business* (Vol. 60) 323-345.

Kothari, S. P. 2001. Capital Markets Research in Accounting. *Journal of Accounting and Economics* (Vol. 31) 105-231.

Lee, Y.-J., K. R. Petroni, and M. I. N. Shen. 2006. Cherry Picking, Disclosure Quality, and Comprehensive Income Reporting Choices: The Case of Property-Liability Insurers. *Contemporary Accounting Research* (Vol. 23) 655-692.

Leisenring, J. J., and L. T. Johnson. 1994. Accounting Research: On the Relevance of Research to Practice. *Accounting Horizons* (Vol. 8, No. 4) 74-79.

Leuz, C., and R. E. Verrecchia. 2000. The Economic Consequences of Increased Disclosure. *Journal of Accounting Research* (Vol. 38) 91-124.

Lev, B., and P. Zarowin. 1999. The Boundaries of Financial Reporting and How to Extend Them. *Journal of Accounting Research* (Vol. 37) 353-385.

Lipe, M. G., and S. E. Salterio. 2000. The Balanced Scorecard: Judgmental Effects of Common and Unique Performance Measures. *Accounting Review* (Vol. 75) 283-298.

Nelson, M. W., J. A. Elliott, and R. L. Tarpley. 2003. How Are Earnings Managed? Examples from Auditors. *Accounting Horizons* (Vol. 17) 17-35.

Schipper, K. 1989. Commentary on Earnings Management. *Accounting Horizons* (Vol. 3) 91-102.

Sloan, R. G. 1996. Do Stock Prices Fully Reflect Information in Accruals and Cash Flows About Future Earnings? *Accounting Review* (Vol. 71) 289-315.

Sterling, R. R. 1973. Accounting Research, Education and Practice. *Journal of Accountancy* (Vol. 136) 44-52.

Wilks, T. J., and M. F. Zimbelman. 2004. Decomposition of Fraud-Risk Assessments and Auditors' Sensitivity to Fraud Cues. *Contemporary Accounting Research* (Vol. 21) 719-745.

Zeff, S. A. 2003. How the U.S. Accounting Profession Got Where It Is Today: Part 1. *Accounting Horizons* (Vol. 17) 189-205.